GECAM卫星快速预处理流程设计与实现*

马福利[1]，陈玲[1]，李冰[2]，郑世界[3]，王平[3]，于勤思[1]
(1.中国科学院国家空间科学中心，北京 100190；2.中国科学院空天信息创新研究院，北京 100094；3.中国科学院高能物理研究所，北京 100049)

摘 要： GECAM是专门针对引力波伽马暴的研究机遇而提出的中国科学院空间科学（二期）先导专项卫星任务。GECAM卫星通过数传、遥测以及北斗短报文三个通道下行触发、事例、并道、工程以及短报文等多种类型的数据，数据预处理过程需要对这些数据进行快速正确处理，以满足科学家对天文事件数据正确性和时效性的要求。本文针对GECAM卫星数据处理时效性要求高、数据连续完整、多信道数据融合处理以及触发事件数据切分等特点，对科学数据处理流程和天文警报信息处理流程进行了设计，概述了关键核心算法的思路与实现方法。卫星发射入轨后，数据预处理软件运行良好，正确性和及时性指标满足要求，为后续爱因斯坦探针（EP）卫星、中法天文卫星（SVOM）等空间天文卫星开展数据预处理工作提供了参考。

关键词： 警报数据；卫星预处理； GECAM卫星；天文；地面段
中图分类号： P171.3 **文献标识码：** A **文章编号：** 1007-2276-(2004)4-0338-05

天文学是一门观测驱动的科学，天文学的重大进展往往源自于新的观测发现。2015 年 9 月 14 日，LIGO 首次成功探测到来自双黑洞合并产生的引力波，推进引力波常规化探测的进程。2017 年 8 月 17 日，人类首次发现双中子星并合产生的引力波及其电磁对应体[1]，直接证实了引力波电磁对应体的存在及其在引力波相关研究中的重要作用，促使引力波天文学研究成为国际热门的研究领域[2]。引力波暴高能电磁对应体全天监测器（Gravitational wave high-energy Electromagnetic Counterpart All-sky Monitor，GECAM）是专门针对引力波伽马暴的研究机遇而提出的中国科学院空间科学（二期）先导专项卫星任务，基于北斗三号导航系统的短报文服务，卫星能够及时下行空间天文警报信息，是我国首个具有即时下行观测数据能力的空间天文望远镜[3]。

GECAM 卫星由有效载荷以及卫星平台组成，有效载荷包括 25 个伽马射线探测器（GRD）、8 个荷电粒子探测器（CPD）、载荷处理器和载荷舱主结构四部分[4][5]，有效载荷将观测的科学数据以及在轨触发事件信息通过卫星平台的数传子系统、测控子系统以及短报文子系统分别下行到地面。数据落地后，数据预处理软件需要对来自多通道下行的数据进行融合处理以及交叉验证，以保证触发数据产品、天文警报数据产品以及事例和并道数据产品的正确性和完整性，处理后生成的各级产品及时发送至科学应用系统的科学家手中用于开展进一步的科学分析。

1 卫星数据预处理特征分析

GECAM 卫星采集的数据包括科学观测数据、载荷工程数据以及平台工程数据，科学观测数据又根据载荷工作模式以及采样率不同分为触发数据、事例数据、时间并道数据和能谱并道数据。卫星平台按照 CCSDS 源包格式[6]，设计不同的应用过程标识符（APID, Application Process Identifier）在大容量存储中对各类数据进行区分标识与存储。在卫星过境期间，按照不同的虚拟信道组织成标准的 AOS 传输帧[7]下行。当数据成功落地后，地面系统需要对数据进行帧校验、帧同步、虚拟信道分离、源包/科学数据包提取与校验、APID 拆分、关键域校验、排序、时间码解算以及物理量转换等操作[8]，处理生成 CDF 和 FITS 格式的 0A/0B/0D 级数据产品。

传统的卫星数传子系统设计一般都是一个 APID 的数据仅通过一个虚拟信道下行[9]，由

*基金项目：中国科学院战略性先导科技专项空间科学（二期）科学卫星地面支撑项目(XDA15040000)资助
收稿日期：2021-04-14； 修订日期：2021-05-13
作者简介：马福利，男，硕士，研究方向：卫星地面数据处理技术、大数据处理与管理技术。 Email:mafuli007@nssc.ac.cn

于 GECAM 卫星对触发数据以及天文警报信息时效性和可靠性要求较高，有效载荷采集的同一个 APID 的事例数据会通过点播数据虚拟信道、暴发数据虚拟信道以及事例数据虚拟信道三个不同的虚拟信道同时下行，增加了地面系统进行融合数据处理以及按事件进行触发数据切分的难度。此外，科学数据分析过程对事例时间高精度要求以及触发数据产品复杂的处理过程也给地面数据预处理软件的设计都带来了挑战。

1) 多备份、多通道数据融合处理

为保证数传信道接收数据的可靠性，避免出现由于星地链路不稳定或者设备的原因导致落地数据不完整的情况发生，地面接收站在进行 GECAM 卫星数据接收时往往安排多个天线备份接收，落地生成多份原始数据。根据处理需求将多份原始数据发送地面预处理软件。地面系统需要对来自测控、北斗短报文以及数传通道下行的多备份数据进行快速预处理，融合各信道天文警报信息，提取触发编号、触发时间、触发位置等信息进行触发数据产品的生产与验证；融合载荷工程数据中的关键参数进行事例数据和并道数据的质量标识等。

2) 事例数据时间解算过程复杂且精度要求高

由于星上存储空间的约束和数据量压缩的需求，在事例数据格式中没有存储数据到达时间的绝对值，每个事例的到达时间采用一个最小分辨为 0.1μs 的本地时间计数器进行记录，物理事例中只记录最低的 24bit，当其向高位进位时产生一个进位事例，记录该计数器的高位信息。为了能够在地面处理过程中确定事例到达的绝对时间，载荷将卫星平台提供的 GPS 秒同步信号的到达时间作为一个 GPS 事例插入到数据流中，另外在每个 GPS 秒同步信号之后，卫星数管会通过 CAN 总线广播其对应的 UTC 时间，载荷也将其作为一个特殊的 UTC 事例插入事例流中。地面系统需要基于有效载荷物理事例、进位事例、GPS 事例以及 UTC 事例共同还原物理事例的到达时间，时间的解算精度要求达到 10E-7s 量级。

3) 基于事件组织的触发数据产品设计

当有效载荷在轨触发软件探测到显著的触发事件后，将通过 CAN 总线发送触发时刻至卫星平台，卫星平台将触发时刻前 50s 和触发时刻后 250s 共计约 300s 左右的事例数据组成触发数据并通过暴发数据虚拟信道下行；卫星平台在接收到触发时刻后，也会组织约 31 条短报文数据包将关键信息下发至地面，这两类信息是触发数据产品的生产的必要输入。触发时刻前后时间段内的数据对科学分析工作非常关键，为方便科学家开展数据产品分析，呈现触发时间段内尽可能完整的数据内容，地面系统需要基于触发事件编号将触发事件时间范围内的科学数据、载荷工作状态参数、轨道、姿态和日月地空间等信息进行统一处理与产品生产。

2 数据预处理流程定义

GECAM 卫星数据预处理过程从接收到卫星的数传原始数据文件开始[10]，采用基于数据驱动的方式，对地面接收站落地的遥测原始文件、数传原始文件以及北斗短报文数据进行批处理，基于 MESOS 统一资源池对批处理作业任务进行并行调度，支持按优先级进行任务调度，保证高优先级的数据得到优先处理。

2.1 科学数据处理流程

事例数据、并道数据和触发数据是有效载荷直接探测到的科学数据，针对这三类科学数据的传输帧层数据处理、源包层数据处理、数据切分、时间解算以及物理量转换等处理过程，我们设计了科学数据处理，生成不同级别和不同种类的数据产品，如图 1 所示。

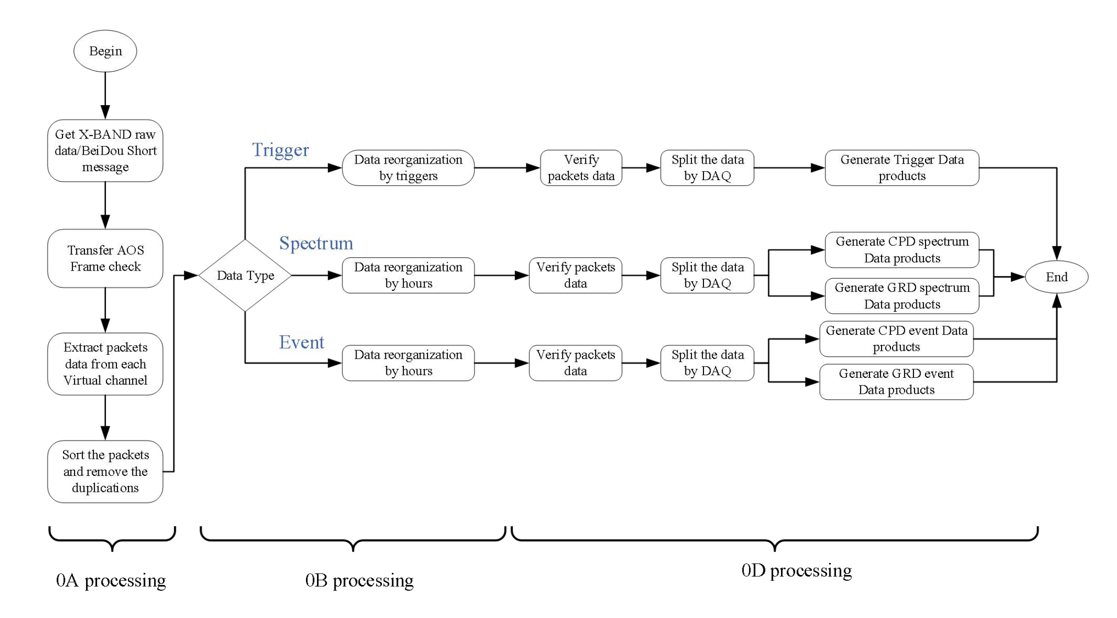


图 1 事例/并道数据处理流程
Fig.1 The processing flow of EVENT and SPECTRUM data

Step-1：AOS 传输帧数据处理，接收并解析数传原始数据，依据帧同步头“1ACFFC1DH”开展 AOS 传输帧同步，对 AOS 进行 LDPC 校验，剔除误码数据，依据虚拟信道标识符对事例数据、并道数据、触发数据和载荷工程数据进行拆分。

Step-2：CCSDS 源包数据处理，提取 AOS 帧数据域内的 CCSDS 源包数据块，依据同步码“146FH”进行 CCSDS 源包定位，根据“包长”提取出源包数据，利用源包序列计数和源包副导头的时间码进行源包排序。输出 GECAM 卫星 0A 级数据产品。

Step-3：数据产品切分，根据 0B 级数据产品格式设计，事例数据和并道数据块需要按小时 hour:00:00~hour:59:59 拆分，基于数据完整性考虑，每个产品将包含前一个小时最后 100 秒的冗余数据；对触发数据产品，按照触发时间进行数据的切分。

Step-4：科学数据包处理，提取 CCSDS 源包数据域中的源数据，进行源数据 CRC 校验，根据源数据类型的不同，分别开展 DAQ、载荷探头的拆分以及源数据到达时间的解算，生成 GRD 和 CPD 载荷的 0D 级事例数据和并道数据。

2.2 天文警报信息处理流程

天文警报信息是指卫星在轨触发软件探测到爆发事件后，按照北斗短报文格式组织的数据，该类信息会分别通过数传信道、遥测信道以及北斗短报文服务进行下行。星上会自主根据触发类型生成不同数量的天文警报信息，当触发类型为长触发时会固定组织生成 31 条天文警报信息，当触发类型为短触发时会组织生成 4≤n≤31 条天文警报信息。

1. 天文警报信息格式定义

天文警报产品按照触发事件编号组织，内容包含一次爆发事件对应的天文警报信息，天文警报信息产品采用 FITS 格式，产品组织如图 2 所示。

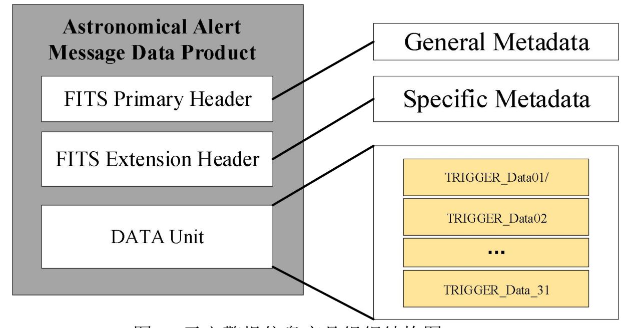


图 2 天文警报信息产品组织结构图
Fig2. The structure of Alert Message data product.

Primary Header 存储产品基本元数据信息，Extension Header 存储数据单元特有的元数据信息，Data Unit 中存储经物理量转换后的参数结果，部分元数据格式如表 1~表 2 所示。

表 1 部分基本元数据信息定义

Table 1 Some basic metadata definition

Key Words	Comments
TELESCOP	Name of the telescope
DATATYPE	Alert Data
TRIGTYPE	Trigger Type
DATE_REF	Reference date for GECAM, expressed in UTC
TSTART	[GECAM MET] Observation start time
TSTOP	[GECAM MET] Observation stop time
SOFTWARE	GECAM data processing software version

表 2 Extension Header 特有的关键元数据信息定义

Table 2 Extension Header dedicated critical metadata definition

Key Words	Comments
PRIORITY	Data downlink priority
FORMCODE	Format code
TRINUM	Trigger number
BDNUM	Trigger sequence number
TRITIME	Trigger time
FTRINUM	Class No. 1 trigger times
PROTRIST	Detector state code
TRIERR	Location error of trigger event
3GRDHTSN	Number of the 3 GRDs that trigger the highest significance
3GRDHTS	Significance of the 3 GRDs that trigger the highest significance
RAJ2000	RA of the trigger event
DECJ2000	DEC of the trigger event
CLASS	Trigger class
RATRIT	RA of the satellite at trigger time in J2000
DECTRIT	DEC of the satellite at trigger time in J2000
TRI_Q1	Trigger attitude Q1 at trigger time + 20s in J2000
TRI_Q2	Trigger attitude Q2 at trigger time + 20s in J2000
TRI_Q3	Trigger attitude Q3 at trigger time + 20s in J2000

2. 警报信息处理流程

天文警报信息最先通过北斗短报文进行下行，为了保证警报信息的准实时性，科学家会优先对该部分信息进行初步的科学分析。本文综合考虑数传信道稳定、数据传输质量较好以及数据内容相对完整等因素，设计了以数传通道天文警报信息为主的数据处理流程，当星地链路问题或者其他原因导致数据内容缺失时，会通过遥测信道以及北斗系统下行的天文警报信息进行融合处理，补充缺失数据。处理流程如图 3 所示。

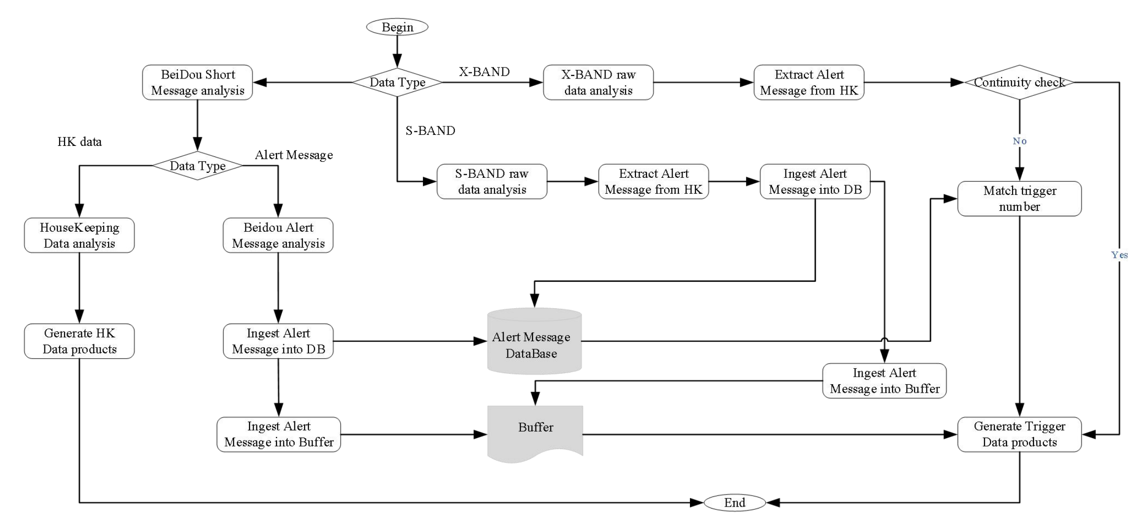


图 3 天文警报信息融合数据处理流程

Fig. 3 The processing flow of alert message

Step-1：对接收的原始数据文件进行类型判断，依据来源不同，采用不同的处理步骤和方法进行天文警报信息的提取。

Step-2：按照格式约定对下行的北斗短报文信息进行解析，判断短报文的数据类型，识别出卫星平台通过短报文下行的警报数据，提取警报事件的位置和时间信息，写入警报信息入库，并将数据写入缓存空间。

Step-3：对遥测信道的数据，直接提取北斗短报文异步包，并进行警报信息入库以及缓存数据的写入。

Step-4：提取来自数传原始数据文件中的平台遥测虚拟信道数据，解析北斗短报文异步包并进行数据连续性的判断。针对数据缺失的情况，读取警报信息库和缓存数据，进行警报信息的匹配以及数据融合，对长触发类型的警报信息进行警报信息完整性的判断与完整性标识，最后生成 FITS 格式的天文警报信息产品。

3 软件设计与实现

3.1 软件实现

本文对 GECAM 卫星数据预处理流程中的各个环节进行了划分，按照数据处理步骤、数据处理级别以及数据类型设计多种 GECAM 卫星数据预处理组件。

1) **0A 级数据处理组件：**对数传和遥测信道中传输帧以及源包进行校验、提取，按照时间码和源包计数排序等处理，输出按照 APID 拆分的不同类型的源包数据。
2) **0B 级事例、并道、载荷工程数据处理组件：**对事例、并道数据源包进行历史数据合并，按照固定时间段进行数据产品切分和完整性校验。
3) **0B 级触发数据处理组件：**对触发数据源包按照触发时间进行切分和组织，按照触发数据产品组织模型进行产品重组，输出完整触发数据产品。
4) **0D 级事例、并道数据、载荷工程数据处理组件：**在 0B 级数据的基础上，对物理事例进行时间解算和能量统计；对各载荷探头的时间和能谱并道统计；对载荷轨道、姿态、太阳月亮星历数据进行解算。
5) **0D 级触发数据处理组件：**在 0B 级触发数据的基础上，对触发时间段的物理事例进行时间解算和能量统计，并匹配该触发事件发生时的载荷轨道、姿态和日月地空间信息数据。
6) **天文警报信息处理组件：**提出数传信道传输的天文警报信息，依据警报信息格式，提取长触发和短触发数据，对北斗短报文处理组件输出信息进行融合处理、交叉验证，输出 FITS 格式的天文警报信息产品。
7) **北斗短报文处理组件：**对北斗下行的北斗短报文数据进行类型判别，根据数据标识识别出平台关键参数和天文警报信息，将天文警报信息存入警报信息库中。
8) **遥测短报文处理组件：**提取遥测信道下行的天文警报信息，将天文警报信息存入警报信息库中。

3.2 关键算法设计

3.2.1 触发数据融合算法

由于空间天文事件具有较高的时效性要求，为了能够及时帮助和引导其他望远镜进行后随观测，GECAM 卫星下行的触发事件数据需要在尽可能短的时间内完成科学分析。触发数据是当星上在轨触发软件发现触发事件时由卫星平台抽取出的该触发时间段的事例数据，是有效载荷探测到的与事件相关的最主要的科学数据，抽取的触发数据段会根据触发事件的顺序进行记录，为了提高触发事件相关分析工作的效率，本文针对基于事件组织的触发数据产品组织模型，设计了触发数据产品融合处理算法。

1) 触发信息提取

对通过北斗短报文的第 1 条和第 2 条短报文进行解析，提取触发编号、触发类型以及触发分类信息存入到触发信息库中。

2) 辅助数据集解析

对工程信道的载荷工程数据进行解析和物理量转换，依据约定的格式和数据类型，提取轨道、姿态、太阳月亮星历以及载荷工作状态数据并输出辅助数据集。

3) 科学数据生成

科学数据包在暴发数据信道中按照触发数据段进行顺序存储，处理时需要依据时间间隔对各个数据段进行区分，提取触发数据中的载荷类型信息和相应的数据段，与触发信息库中的事件基本信息进行匹配，输出按照触发编号、载荷类型以及触发类型分类的触发数据集。

4) 产品融合组织

最后按照触发事件进行触发数据集和辅助数据集的融合与重组，输出针对每个特定触发事件的触发数据产品。

3.2.2 高精度事例时间解算算法

本文提出了高精度事例时间解算算法，该算法在对本底物理事例和 GPS 事例的进位信息迭代和融合的基础上，通过对 GPS 事例信息的高阶拟合，实现对单一物理事例的高精度时间解算，算法具体包括以下 4 个模块：

1) 基于本底物理事例的进位信息迭代

本底物理事例在正常时间范围内（不包括 SAA 区和偏压状态下）都是均匀生成，其晶振低位计数会出现规律性翻转。利用这一特性，对本底物理事例的低位计数进行翻转次数迭代，能够获得每一个物理事例的高位计数，生成基础进位信息。

2) 基于 GPS 事例的进位信息迭代

根据 GPS 事例的生成频率和高位进位频率的耦合性，能够获得每一个 GPS 的高位计数。通过相邻两个 GPS 事例的低位信息，通过高位进位频率的适应性迭代，能够获得相邻两个 GPS 事例相隔的进位信息。对连续 GPS 事例的进位信息进行迭代处理，生成基于 GPS 事例的进位信息。

3) 双重进位信息融合

根据 GPS 事例和物理事例的生成规则，确定基于物理事例的进位信息和基于 GPS 事例的进位信息的绝对差，完成对 GPS 的进位信息进行校正。寻找特定的 GPS 事例，该事例的前后物理事例的进位信息均相同，将该 GPS 事例的进位信息和前后物理事例的进位信息进行差值运算，获取绝对差。对所有 GPS 事例的进位信息进行绝对差修正，完成双重进位信息的融合匹配。

4) 基于 GPS 信息的高阶拟合

在获取 GPS 事例的进位信息的基础上，通过对 GPS 事例与 UTC 事例的一一对应关系，获取 GPS 事例的绝对时间。对 GPS 事例的进位信息和低位计数进行计算，获取 GPS 事例的完整晶振计数。通过 GPS 事例的绝对时间-完整晶振计数的对应关系，进行高阶拟合后，获取高精度的晶振频率以及相应的处理系数。在获取晶振频率后，对物理事例的完整晶振计数进行拟合函数运算，获得每一个物理事例的高精度时间。

4 应用验证

目前该数据预处理流程设计以及算法实现已经应用在 GECAM 卫星地面支撑系统数据处理与产品生产软件中，支持了 GECAM 卫星的在轨测试。GECAM 任务每轨下行约 15GB 左右的科学数据，每次下行数据包括工程数据、事例数据和并道数据，触发数据的数据量不固定。数据处理时效性如图 4 所示，运行结果表明，本文实现的科学数据处理流程和天文警报信息处理流程满足 GECAM 卫星数据预处理时效性的要求。经时间解算后得到的各时刻的 GRD 光子计数、CPD 荷电粒子计数均和时间并道统计的结果一致，误差在 10e-7 以内，如图 5 所示，满足科学数据分析对时间精度的要求。

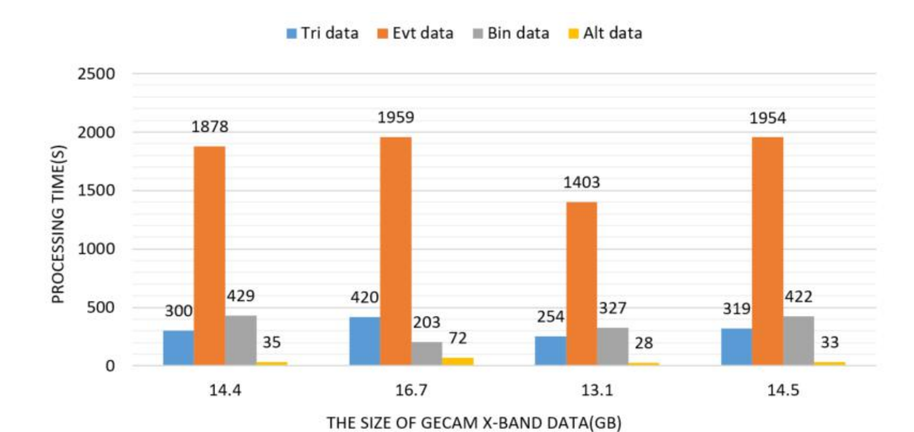
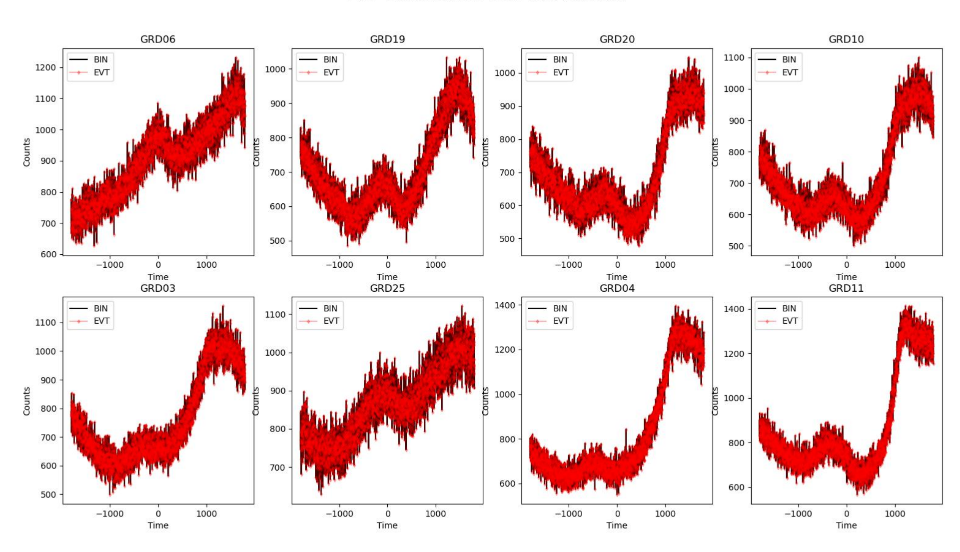
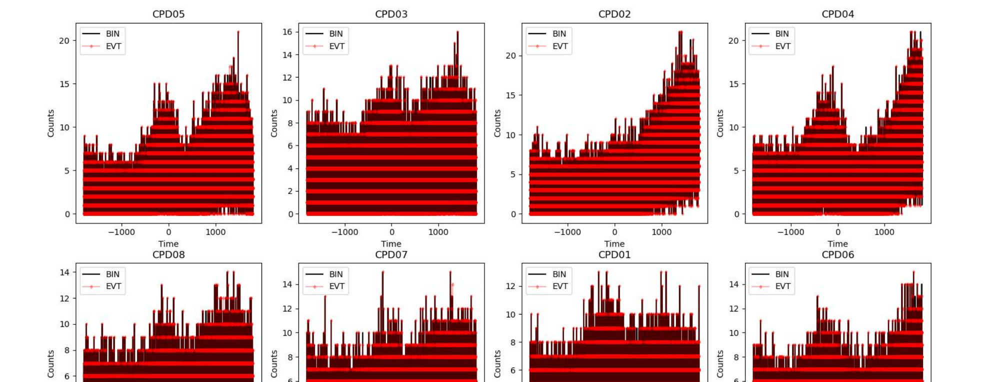


图 4GECAM 数传原始数据预处理时间
Fig.4 The proessing time of GECAM X-band data



（a）2021-03-24 02:28:20±1800 秒内 GRD 物理事例统计的光子计数和时间并道统计的光子计数对比图
(a) The comparison of statistic photon counts between GRD EVENTS and SPECTRUM data in 2021-03-24 02:28:20±1800s



（b）2021-03-24 02:28:20±1800 秒内 CPD 物理事例统计的荷电粒子计数和时间并道统计的荷电粒子计数对比图
(b) The comparison of statistic particle counts between GRD EVENTS and SPECTRUM data in 2021-03-24 02:28:20±1800s
图 5 物理事例和时间并道计数统计对比图

Fig 5 The comparison between EVENTS and SPECTRUM statistic data

5 总结与展望

GECAM 卫星我国首个具有即时下行观测数据能力的空间天文望远镜，本文针对 GECAM 卫星科学数据处理需求，设计了科学数据处理流程和天文警报信息处理流程，解决了触发数据融合算法以及高精度事例数据时间解算算法等关键问题，研制了 GECAM 快速预处理软件。经过在轨测试期间的试运行，验证了快速预处理软件的功能和性能指标。本文提出的基于事件组织的触发数据产品组织模型以及高精度事例时间解算算法等内容，对我国后续的 EP、SVOM 等卫星任务的数据产品的设计与预处理系统的研制具有良好的参考意义。

致谢：感谢国家科技资源共享服务平台-国家空间科学数据中心（http://www.nssdc.ac.cn）提供数据资源/数据分析环境/计算服务/应用平台支持。

We acknowledge the National Space Science Data Center of China for providing data resources/ computing services/ data analysis environment.

参考文献

[1] ABBOTT B P, ABBOTT R, ABBOTT T D, et al. GW170817: observation of gravitational waves from a binary neutron star inspiral[J]. Physical Review Letters, 2017, 119(16): 161101.

[2] Li T P. et al. (The Insight-HXMT team). Insight-HXMT observations of the first binary neutron star merger GW170817. Sci China-Phys MechAstron, 2018, 61: 031011.

[3] 北斗三号卫星导航系统助力宇宙探测GECAM卫星准实时下传伽马暴观测警报[N]. http://www.ihep.cas.cn/xwdt/gnxw/2020/202101/t20210121_5874777.html.

[4] 韩兴博，张科科，黄佳等，GECAM 双星总体方案及关键技术[J]，中国科学：物理学力学天文学，2020, 50(12):129507.

Han Xingbo, Zhang Keke, Huang Jia, et al. GECAM satellite system design and technological characteristic [J]. Scientia sinica physica, Mechanica & astronomica, 2020, 50(12):129507.

[5] 李新乔，文向阳，安正华，等. GECAM 卫星有效载荷介绍. [J] 中国科学: 物理学力学 天文学, 2020, 50(12): 129508.

Li xinqiao, Wen xiangyang, An Zhenghua, et al. The GECAM and its payload [J]. Scientia sinica physica, Mechanica & astronomica, 2020, 50(12): 129508.

[6] Space packet protocol: CCSDS 133.0-B-1, Blue Book [S], Washington DC：2003-09.

[7] AOS space data link protocol, CCSDS 732.0-B-2 [S], Washington DC: 2006-07.

[8] 赵海升，葛明玉，李正恒等. 一种天文卫星数据预处理方法 [J]. 天文研究与技术，2017,14（3）：376-381.

Zhao Haisheng, Ge Mingyu, Li Zhengheng, et al. A method of Data Preprocessing for Astronomical Satellite [J]. Astronomical Research &Technology, 2017, 14(3):376-381.

[9] 赵海升，樊凤霞，贾淑梅等. 一种卫星下传数据解析方法[J]. 天文研究与技术，2019,16（4）：517-522.

Zhao Haisheng, Fan fengxia, Jia Shumei, et al. A Decoding Method of Data Packets for Satellite Data [J]. Astronomical Research &Technology, 2019,16（4）：517-522.

[10] 陈维，宋黎明，郑世界等. GECAM 卫星科学应用系统简介[J]. 中国科学：物理学力学天文学，2020，50(12)：129512.

Chen wei, Song Liming, Zheng Shijie, et al. Introduction of the scientific application system of GECAM [J]. Scientia sinica physica, Mechanica & astronomica, 2020 , 50(12)：129512.

Design and Implementation of GECAM Preprocessing pipeline*

Ma Fuli[1], Chen Ling [1], Li Bing [2], Zheng Shijie[3], Wang Ping[3], Ge Mingyu[3], Yu Qinsi[1]

(1. National Space Science Center, Chinese Academy of Sciences, Beijing 100190; 2. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094; 3. Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049)

Abstract: The Gravitational wave high-energy Electromagnetic Counterpart All-sky Monitor (GECAM) is a space science project proposed by the Chinese Academy of Sciences specifically for exploring the significant opportunity of the gravitational wave multi-messenger astronomy by monitoring the gravitational wave high energy electromagnetic counterpart in all-sky. GECAM satellite can downlink the EVENT data, housekeeping data, and alarm data through X-band, S-band and also Beidou-3. The ground segment preprocessing process needs to process these data quickly and correctly to meet scientists requirement for correctness and timelines. In this paper, we designs the scientific data processing flow and the astronomical warning information processing flow according to the characteristics of GECAM data processing, and also we outlines the key core algorithms and shows the implementation of the key methods. After the satellite was successfully launched, the data preprocessing software runs well, and the accuracy meet the system requirements. The work done in this paper provides guidelines for the follow-up Einstein Probe (EP) satellite, Sino-French Astronomical Satellite (SVOM) and other space astronomical satellites.

Keywords: Alert Message; Satellite data preprocessing; GECAM satellite; Astronomy; Ground segment